



#4

PATENT
3626-0240P

IN THE U.S. PATENT AND TRADEMARK OFFICE

Applicant: Watson WU Conf.: 8731
Appl. No.: 10/034,390 Group:
Filed: January 3, 2002 Examiner:
For: METHOD AND SYSTEM FOR PRODUCING A BOOK
FROM A VIDEO SOURCE

LETTER

Assistant Commissioner for Patents
Washington, DC 20231

March 12, 2002

Sir:

Under the provisions of 35 U.S.C. § 119 and 37 C.F.R. § 1.55(a), the applicant(s) hereby claim(s) the right of priority based on the following application(s):

<u>Country</u>	<u>Application No.</u>	<u>Filed</u>
TAIWAN	90122705	September 13, 2001

A certified copy of the above-noted application(s) is(are) attached hereto.

If necessary, the Commissioner is hereby authorized in this, concurrent, and future replies, to charge payment or credit any overpayment to Deposit Account No. 02-2448 for any additional fee required under 37 C.F.R. §§ 1.16 or 1.17; particularly, extension of time fees.

Respectfully submitted,

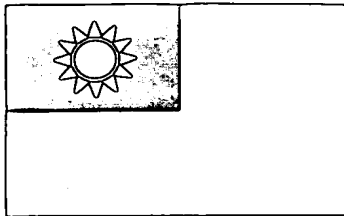
BIRCH, STEWART, KOLASCH & BIRCH, LLP

By Joe McKinney Muncy
Joe McKinney Muncy, #32,334

KM/ndb
3626-0240P

P.O. Box 747
Falls Church, VA 22040-0747
(703) 205-8000

Attachment



Watson WU
3626-0240P
101034, 390
January 3, 2002
BSKB, LLP
(703) 205-8000



中華民國經濟部智慧財產局

INTELLECTUAL PROPERTY OFFICE
MINISTRY OF ECONOMIC AFFAIRS
REPUBLIC OF CHINA

茲證明所附文件，係本局存檔中原申請案的副本，正確無訛，
其申請資料如下：

This is to certify that annexed is a true copy from the records of this
office of the application as originally filed which is identified hereunder:

申請 日：西元 2001 年 09 月 13 日
Application Date

申請 案 號：090122705
Application No.

申請 人：力新國際科技股份有限公司
Applicant(s)

局 長
Director General

陳 明 邦

發文日期：西元 2002 年 1 月 22 日
Issue Date

發文字號：09111001167
Serial No.

申請日期：	案號：
類別：	

(以上各欄由本局填註)

發明專利說明書

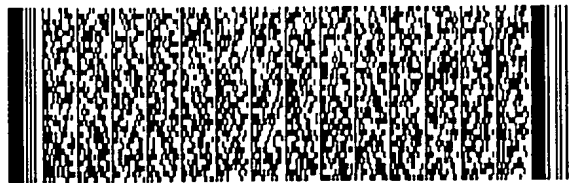
一、 發明名稱	中 文	圖書製作系統與方法
	英 文	
二、 發明人	姓 名 (中文)	1. 吳昌隆
	姓 名 (英文)	1. Watson Wu
	國 籍	1. 中華民國
	住、居所	1. 30077 新竹科學工業園區園區二路9-1號
三、 申請人	姓 名 (名稱) (中文)	1. 力新國際科技股份有限公司
	姓 名 (名稱) (英文)	1. NewSoft Technology Corporation
	國 籍	1. 中華民國
	住、居所 (事務所)	1. 11550 台北市南港軟體工業園區三重路19-3號3樓
	代表人 姓 名 (中文)	1. 黃崇仁
	代表人 姓 名 (英文)	1.



四、中文發明摘要 (發明之名稱：圖書製作系統與方法)

本發明提供一種圖書製作系統，其係用來產生包括一文字部分與一插圖部分之一圖書，且包括一視訊接收模組、一解碼模組、一文字擷取模組、一插圖擷取模組以及一圖書產生模組。在本發明中，視訊接收模組接收一原視訊資料；解碼模組將原視訊資料解碼成為一視訊資料；文字擷取模組則依據一製作方針自視訊資料中取得文字部分；插圖擷取模組則依據製作方針自視訊資料中擷取至少一關鍵畫面 (key frame) 以作為插圖部分；圖書產生模組將所取得之文字部分與插圖部分排版以產生圖書。本發明亦揭露一種依據上述系統實施的圖書製作方法。

英文發明摘要 (發明之名稱：)



本案已向

國(地區)申請專利

申請日期

案號

主張優先權

無

有關微生物已寄存於

寄存日期

寄存號碼

無

五、發明說明 (1)

【發明領域】

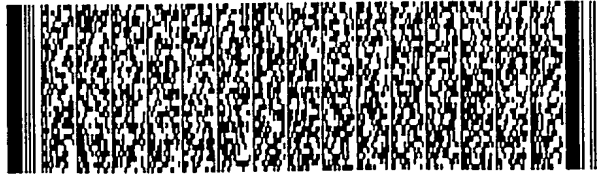
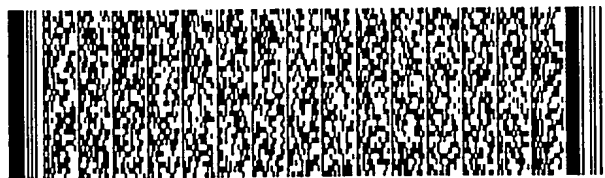
本發明係關於一種圖書製作系統與方法，特別關於一種利用一電腦軟體分析一視訊源 (Video) 以自動產生繪本、畫冊、漫畫、電子書等圖書文件的圖書製作系統與方法。

【習知技術】

根據目前的技術，一般在製作繪本、畫冊、漫畫、電子書等圖書時，其內容的來源通常仍利用人工繪圖，或是藉由電腦針對單張影像一一編輯整理，以便彙編成書冊。

然而，隨著數位攝影機、電視卡 (TV Tuner Card)、機上盒 (Setup Box)、DVD、VCD 等電子資訊產物的日益普及，使用者可以很容易地取得數位視訊，因此，利用電腦處理視訊源以產生圖書文件，儼然成為電腦多媒體領域的重要應用與需求。

如上所述，當所得到的影像資料不是單張影像而是連續影像的視訊源時，使用者必須將連續影像的視訊源分解成複數張影像，然後才能夠藉由電腦針對該等影像編輯整理成冊。然而，對於一般的視訊內容 (Video Content) 而言，在NTSC標準中，其撥放一秒鐘可能是包含29.97張影像的連續切換，而在PAL標準中，其撥放一秒鐘可能是包含25張影像的連續切換，所以，一分鐘長的視訊內容便具有1500~1800張影像，如果使用者一一編輯每一張影像，將是一件非常耗時而沒有效率的事。



五、發明說明 (2)

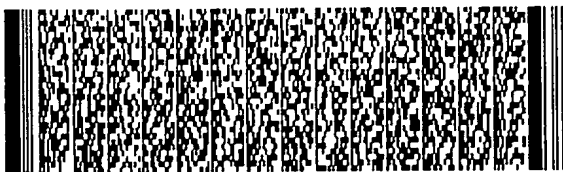
因此，如何能夠有效率地利用視訊內容來產生繪本、畫冊、漫畫、電子書等圖書文件，正是當前一個重要的課題。

【發明概要】

針對上述問題，本發明之目的為提供一種圖書製作系統與方法，其係能夠自動分析一視訊源以產生繪本、畫冊、漫畫、電子書等圖書文件。

為達上述目的，依本發明之圖書製作系統係用來產生包括一文字部分及一插圖部分的一圖書，且包括一視訊接收模組、一解碼模組、一文字擷取模組、一插圖擷取模組以及一圖書產生模組。在本發明中，視訊接收模組接收一原視訊資料，解碼模組解碼原視訊資料以取得一視訊資料，而原視訊資料可以是任意一種視訊格式，文字擷取模組則依據一製作方針自視訊資料中取得文字部分，插圖擷取模組則依據製作方針自視訊資料中擷取至少一關鍵畫面 (key frame) 以作為插圖部分，然後圖書產生模組依據所取得之文字部分與插圖部分產生圖書。

另外，依本發明之圖書製作系統更包括一編輯模組、一圖書格式 (template) 選擇模組、以及一製作方針選擇模組。在本發明中，製作方針選擇模組接受一使用者選擇所需之製作方針，編輯模組接收使用者之操作以便對圖書之內容進行編輯，圖書格式選擇模組接收使用者選擇所需的至少一圖書格式，而圖書產生模組便套用所選出之圖書



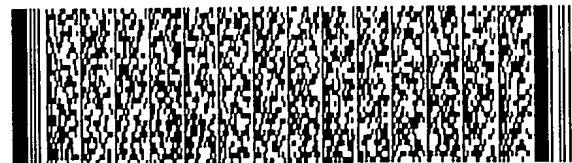
五、發明說明 (3)

格式來排版文字部分與插圖部分以產生圖書。

如上所述，製作方針選擇模組所能夠選擇之製作方針係包括一音訊 (audio) 分析演算法則、一字幕 (caption) 分析演算法則、一場景/鏡頭變換分析演算法則以及一影像分析演算法則，其中，音訊分析演算法則係一種分析視訊資料之音訊資料的演算法則；字幕分析演算法則係一種分析視訊資料之字幕資料的演算法則；場景/鏡頭變換分析演算法則係一種分析視訊資料之場景/鏡頭變換資料的演算法則；影像分析演算法則係一種分析視訊資料之影像資料的演算法則，而且其可以將影像資料與預先提供之一影像範例資料作比對分析，或是將影像資料與預先提供之一物體資料作比對分析，或是分析出影像資料中的一字幕影像資料。

因此，文字擷取模組與插圖擷取模組能夠依據上述之音訊分析演算法則、字幕分析演算法則、場景/鏡頭變換分析演算法則、或是影像分析演算法則來取得製作圖書所需的文字部分與插圖部分等資料，接著，圖書產生模組將上述文字部分與插圖部分套入圖書格式中，於是便自動產生繪本、畫冊、漫畫、電子書等圖書文件。

本發明亦提供一種圖書製作方法，其包括一視訊接收步驟、一解碼步驟、一文字擷取步驟、一插圖擷取步驟以及一圖書產生步驟。在本發明中，視訊接收步驟先接收原視訊資料，接著解碼步驟解碼原視訊資料以取得視訊資料，然後文字擷取步驟與插圖擷取步驟分別自視訊資料中



五、發明說明 (4)

擷取出製作圖書所需之文字部分與插圖部分，最後圖書產生步驟依據文字部分與插圖部分產生圖書。

另外，依本發明之圖書製作方法更包括一編輯步驟以於圖書產生後編輯圖書之內容、一圖書格式選擇步驟以便使用者選取所需之圖書格式，進而讓圖書產生步驟套用該圖書格式來產生圖書、以及一製作方針選擇模組以便使用者選取所需之製作方針。

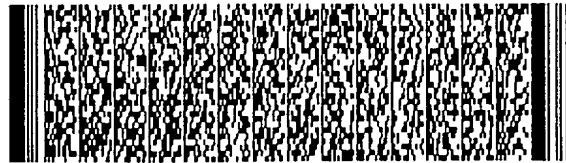
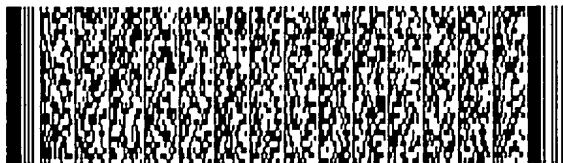
由於依本發明之圖書製作系統與方法能夠自動分析一視訊源，並配合多種視訊格式，且整合視訊內容分析、文字辨識、聲音辨識等技術，來產生繪本、畫冊、漫畫、電子書等圖書文件，所以能夠有效率地利用視訊內容來產生圖書文件。

【較佳實施例之詳細說明】

以下將參照相關圖式，說明依本發明較佳實施例之圖書製作系統與方法，其中相同的元件將以相同的參照符號加以說明。

請參照圖1所示，依本發明較佳實施例之圖書製作系統1係用來產生包括一文字部分801以及一插圖部分802之一圖書80，且包括一視訊接收模組101、一解碼模組102、一製作方針選擇模組103、一文字擷取模組104、一插圖擷取模組105、一圖書格式選擇模組106、一圖書產生模組107以及一編輯模組108。

在本實施例中，圖書製作系統1可以應用於一電腦設



五、發明說明 (5)

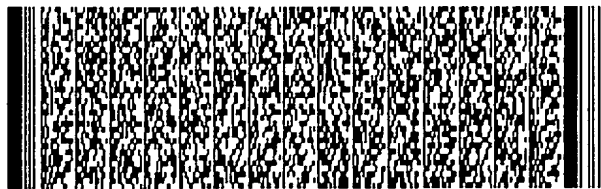
備60中，而電腦設備60可以是習知的電腦裝置，其包括一訊號源介面601、一記憶體602、一中央處理單元(CPU)603、一輸入裝置604以及一儲存裝置605。其中，訊號源介面601係與一訊號源輸出裝置或是一訊號源紀錄裝置連接，例如是光碟機、FireWire (IEEE 1394 Interface)、通用序列埠(USB)等介面裝置，而訊號源輸出裝置例如是數位攝影機，訊號源紀錄裝置例如是VCD、DVD等。記憶體602可以是DRAM或EEPROM等任何一種或數種設置於電腦裝置中的暫存記憶體。中央處理單元603則可採用任何一種習知之中央處理器架構，例如，包括ALU、暫存器與控制器等，以進行各種資料之處理與運算，以及控制電腦設備60中各元件的作動。輸入裝置604可以是滑鼠、鍵盤等能夠由使用者自行輸入訊息，或是操作各軟體模組的裝置。儲存裝置605可以是硬碟機、軟碟機等任何一種或數種電腦可讀取之資料儲存裝置。

本實施例中的各模組係指儲存於儲存裝置605中或是一紀錄媒體中的軟體模組。中央處理單元603於讀取各模組之後，即可經由電腦設備60中的各元件來實現各模組的功能。然而需注意者，熟習該項技術者亦可將本實施例中所揭露之軟體模組製作成硬體，如ASIC

(application-specific integrated circuit) 晶片等，而不違反本發明之精神與範疇。

以下詳細說明本實施例中各模組之功能。

在本實施例中，視訊接收模組101接收一原視訊資料

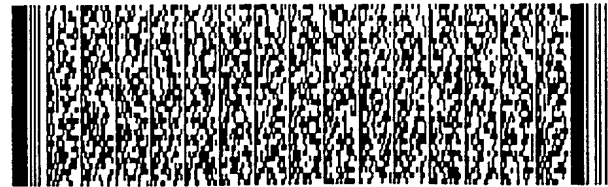
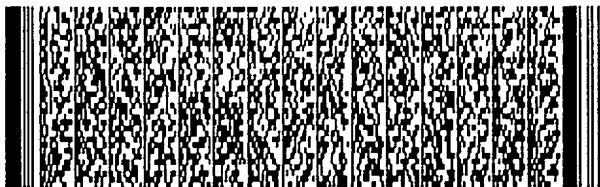


五、發明說明 (6)

40，解碼模組102解碼原視訊資料40以取得一視訊資料41，製作方針選擇模組103係接受一使用者的操作以選取所需之一製作方針50，文字擷取模組104則依據製作方針50自視訊資料41中取得文字部分801，插圖擷取模組105則依據製作方針50自視訊資料41中擷取至少一關鍵畫面以作為插圖部分802，而圖書格式選擇模組106接收使用者之選擇以提供至少一圖書格式70，圖書產生模組107係套用圖書格式70，並依據所取得之文字部分801與插圖部分802產生圖書80，最後，編輯模組108於圖書80產生之後，接受使用者操作以編輯圖書80之內容。

如上所述，視訊接收模組101係與訊號源介面601配合，例如，視訊接收模組101可以透過FireWire(IEEE 1394 Interface)取得儲存於數位攝影機中的原視訊資料40，或是透過光碟機取得記錄於VCD、DVD中的原視訊資料40。原視訊資料40係由各種視訊擷取裝置或接收裝置如數位攝影機、電視卡、機上盒等，以及各種視訊儲存裝置如DVD、VCD所儲存、傳送、廣播(Broadcasting)或接收的視訊源，且其能夠以各種視訊資料格式(如MPEG-1, MPEG-2, MPEG-4, AVI, ASF, MOV等)儲存、傳送、廣播或接收。

解碼模組102能夠針對輸入的原視訊資料40之視訊格式、編碼方式、或壓縮方式進行解碼轉換還原為編碼前之資料或近似於編碼前之資料，例如，若編碼方式採用失真壓縮方式(Lossy Compression)，則解碼後只能夠取得近似於編碼前之資料，以便產生一視訊資料41。在本實施例

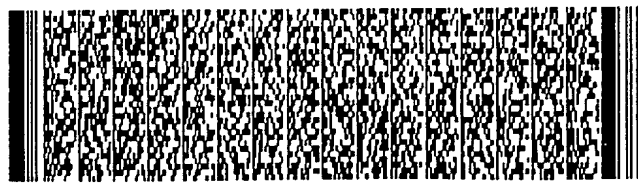
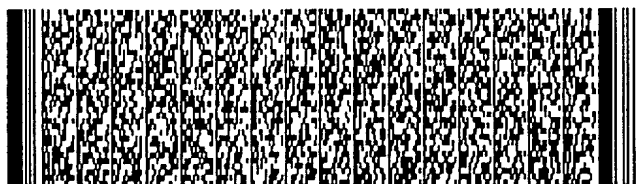


五、發明說明 (7)

中，視訊資料41包括一音訊資料411、一字幕資料412以及一影像資料413。音訊資料411為視訊資料41中所撥放的聲音；字幕資料412為配合影像資料413出現於螢幕上的字幕串流(caption stream)；影像資料413為視訊資料41所顯示的所有單張影像，通常每秒鐘的視訊資料41係由25張單張影像或29.97張單張影像連續撥放所構成。

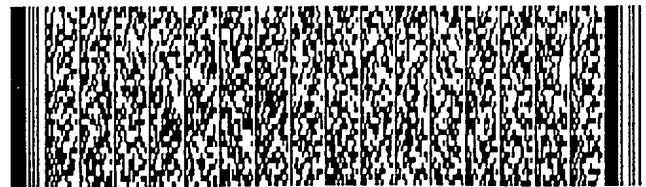
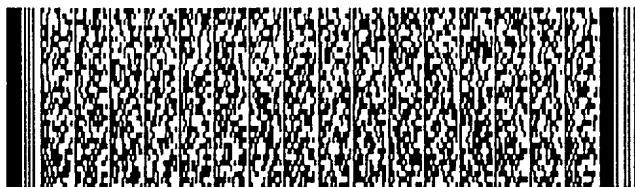
製作方針選擇模組103係與輸入裝置604配合，以便由使用者利用輸入裝置604選擇製作圖書80時所必須遵循的方針，而依本實施例所提供的製作方針50包括一音訊分析演算法則501、一字幕分析演算法則502、一影像分析演算法則503以及一場景/鏡頭變換分析演算法則504。

承上所述，音訊分析演算法則501係分析視訊資料41的音訊資料411，並利用特徵抽取(Features Extraction)與特徵匹配(Features Matching)方式進行分析。音訊資料411之特徵包括如頻譜特徵(Spectral Features)、音量(Volume)、零軸交會率(Zero Crossing Rate)、音調(Pitch)等。如上所述，當抽取頻譜特徵(Spectral Features)後，其經由雜音衰減(Noise Reduction)、分段(Segmentation)，並利用快速傅利葉轉換(Fast Fourier Transform)將音訊資料411轉至頻率域(Frequency)，然後由一組頻率濾波器(Filters)進行特徵值抽取，這組特徵值組成一個頻譜特徵向量(Spectral Feature Vector)。音量是容易量測之一種特徵，其可利用均方根值(RMS, Root Mean Square)代表其特徵值，然後藉由音量



五、發明說明 (8)

(Volume) 分析可輔助分段(Segmentation)之進行，亦即透過靜音偵測(Silence Detection)幫助音訊資料411段落邊界(Boundaries)之決定。零軸交會率(Zero Crossing Rate)為計算每段(Clips)聲音波形(Waveform)與零軸(Zero Axis)交會之次數。音調(Pitch)為聲音波形(Waveform)的基頻(Fundamental Frequency)。因此，音訊資料411可利用上述之音訊特徵及其特徵值所組成之特徵向量(Feature Vector)與音訊樣本(Audio templates)之特徵進行分析比對，以便取得所需的音訊資料411，並經由語音辨識技術取得文字部分801，並取得於視訊資料41中與所需之音訊資料411對應之影像資料413以作為插圖部分802。在本實施例中，音訊分析演算法則501可以預先提供音訊樣本類別，如音樂(Music)、語音(Speech)、動物聲(Animal Sound)、男聲(Male Speech)與女聲(Female Speech)等，以供使用者選擇所欲尋找之音訊類別，因此，特徵匹配便於容許的距離範圍內，尋找與音訊資料411的特徵向量(Feature Vector)具有最短幾何距離(Euclidean Distance)的音訊樣本類別，若此最接近之音訊樣本類別與使用者所選擇之音訊類別相同，則該音訊資料411符合搜尋條件，另外，可以利用最短幾何距離的倒數(Inverse)來表示所選擇之音訊資料411的可信度(Confidence)，從符合音訊樣本類別之音訊資料411找出對應之視訊畫面段落(Clips)，並從這些視訊畫面段落的每一鏡頭中挑選出符合取圖需求的影像作為插圖部分



五、發明說明 (9)

802。另外，若視訊資料41包括字幕串流(Caption Stream)，則解讀所選擇之音訊資料411所對應之視訊資料41內的字幕串流，來作為圖書80之文字部分801；若視訊資料41未包括字幕串流，則解讀所選擇之音訊資料411內之音訊資料411並利用語音分析(Speech Analysis)進行語音與文字(Voice to Text)的轉換處理，以作為圖書80之文字部分801。另外，音訊分析演算法則501之運算複雜度低於影像或視覺(Visual)分析，並可作為影像或視覺(Visual)分析之引導及輔助資料。

另外，字幕分析演算法則502係分析視訊資料41中的字幕資料412，並篩選具有字幕之視訊畫面。換言之，若視訊資料41包括字幕串流則解讀字幕串流以作為文字部分801，並尋找與字幕對應且時間同步之第一個視訊畫面作為插圖部分802；若視訊資料41未包括字幕串流，而是字幕包含於視訊影像中則利用文字辨識技術將字幕(Caption)從視訊影像中抽取出來作為文字部分801，並針對篩選取得之視訊影像進行影像處理移除字幕(可藉由前後視訊影像之資料進行運算處理)，以取得無字幕的視訊影像以作為插圖部分802。如上所述，文字辨識技術主要利用光學文字辨識技術(OCR, Optical Character Recognition)進行文字辨識。

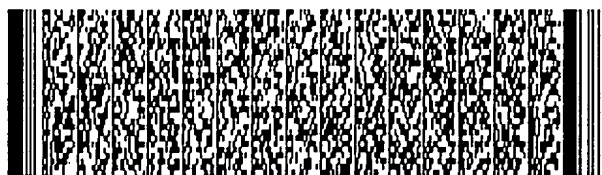
影像分析演算法則503係分析視訊資料41中的一影像資料413，並以色彩、紋理、形狀、動作、位置等基本視覺特徵為分析判斷之依據。在本實施例中，當字幕包含於



五、發明說明 (10)

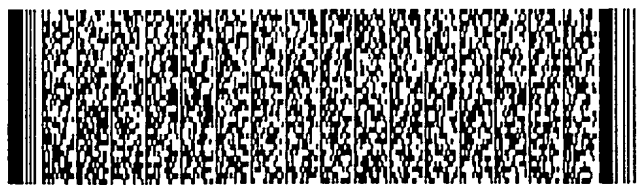
視訊影像時，利用文字辨識技術將字幕從視訊影像中抽取出來作為文字部分801；另外，將視訊資料41與一影像範例資料5031做比較，以便找尋影像視覺特徵相似性大之畫面，或是找尋影像視覺特徵相異性大之畫面以作為插圖部分802，或是將視訊資料41與一物體資料5032作比對，例如以臉部偵測(Face Detection)技術尋找視訊資料41中具有人臉的視訊畫面以作為插圖部分802。在本實施例中，當挑選與影像範例資料5031或是物體資料5032之視覺特徵相似性大之畫面，或是影像視覺特徵相異性大之視訊畫面，以作為符合篩選畫面準則之視訊資料41時，同一鏡頭可設定僅篩選一個畫面以作為插圖部分802。

場景/鏡頭變換分析演算法則504係分析視訊資料41中影像資料413的場景/鏡頭變換，並篩選視訊資料41中影像資料413的場景/鏡頭變換後第一個符合條件之畫面，以作為圖書80之插圖部分802以及視訊資料41之段落的分割點。亦即是，若視訊資料41包括字幕串流則解讀視訊資料41之段落內的字幕資料412以作為圖書80之文字部分801；若視訊資料41未包括字幕串流則解讀視訊資料41之段落內的音訊資料411，並利用語音分析進行語音與文字的轉換處理以作為圖書80之文字部分801。一般而言，視訊資料41係為一視訊串列(Video Sequence)，其通常由許多場景(Scenes)所組成，而每一場景又由複數個鏡頭(Shots)所組成。在影片中，其最小單位是一個鏡頭，而影片便是由許多的鏡頭所堆砌起來的；在劇本中，其最小單位是一個



五、發明說明 (11)

場景，或稱作場戲，場景表示每一故事或題材之段落，每一場景具有一明確的事件發生起始點，也具有一明確的結束點，在這樣的一段時間範疇中便稱作一場景，或稱作一場戲。通常，一個鏡頭由複數個視覺特性(如色彩(Color)、紋理(Texture)、形狀(Shape)、動作(Motion))具一致性之畫面(Frames)所組成，並且，其係依據攝影機運動方向(Camera Direction)與攝影取景角度(View Angle)的改變而有變化，例如，當攝影機以不同之攝影取景角度來拍攝同一場景時，會產生不同之鏡頭，或以相同之攝影取景角度但拍攝不同之區域時，亦會產生不同之鏡頭。由於鏡頭可由一些基本視覺特性而區分，因此將視訊資料41分割成複數個連續的鏡頭是相當容易達成的，此技術主要藉由分析一些基本視覺特性之統計資料如視覺特性柱狀圖(Histogram)，因此，當一畫面之視覺特性與前一畫面之視覺特性差異達到某一程度時，就可在此畫面與前一畫面間作一分割，此分鏡技術亦廣泛運用於視訊編輯軟體。如上所述，將連續具關聯性之鏡頭聚成一場景係為場景變換分析之目的，嚴謹的說，其必須了解視訊資料41之語意及內容，不過結合音訊與視覺特性之分析亦可達到相當程度合理之場景變換分析，通常場景變換會同時產生音訊特性(如音樂、語音、雜音(Noise)、靜音(Silence))與視覺特性(如色彩、動作)之性質變化，而鏡頭之分割只針對視覺特性進行分析，場景變換分析須同時倚重音訊特性與視覺特性之分析。



五、發明說明 (12)

文字擷取模組104與插圖擷取模組105可以是儲存在儲存裝置605之一軟體模組，並透過中央處理單元603之運算來依據製作方針50擷取所需之文字部分801與插圖部分802，以作為製作圖書80之內容。

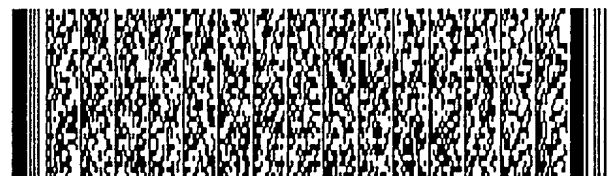
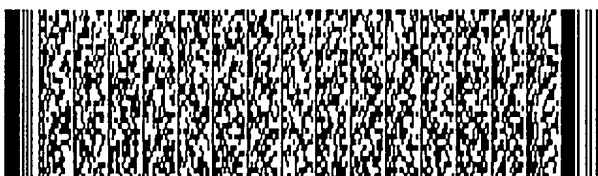
圖書格式選擇模組106所提供的圖書格式70係如繪本、畫冊、電子書、漫畫等，並且可以配合不同之濾鏡(Filters)如藝術家式濾鏡(Artistic Filters)、素描濾鏡(Sketch Filters)、邊緣濾鏡(Edge Filters)，來套用所取得之插圖部分802，以得到使用者想要之影像處理效果(Effects)，而圖書格式70與各種濾鏡係儲存於儲存裝置605中。

圖書產生模組107可以是儲存在儲存裝置605之一軟體模組，並透過中央處理單元603之運算，以便套用圖書格式70，並利用如調整大小(Rescaling)、影像合成(Image Composing)、製作圖框等影像處理(Image processing)功能，來處理所取得之文字部分801與插圖部分802，以便配合使用者選擇之圖書格式70與字型、大小來產生圖書80。

最後，編輯模組108可以與輸入裝置604配合，以便使用者於圖書80產生之後，利用輸入裝置604之操作來進一步編輯圖書80之內容。

為使本發明之內容更容易理解，以下將舉一實例，以說明依本發明較佳實施例之圖書製作方法的流程。

請參照圖2所示，在依本發明較佳實施例之圖書製作方法2中，步驟201係接收原視訊資料40，例如，可以將數

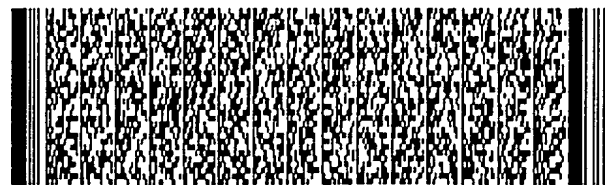
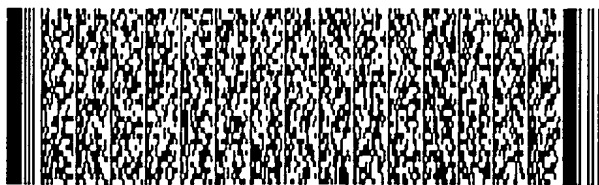


五、發明說明 (13)

位攝影機中紀錄的資料經由傳輸線送至訊號源介面601，以提供作為製作圖書80的畫面與內容。

在步驟202中，解碼模組102係辨識原視訊資料40之格式並解碼原視訊資料40以產生經過解碼之視訊資料41，例如，原視訊資料40為Interlaced MPEG-2格式，亦即是，一個訊框係由兩個訊場(field)所組成，所以，在此步驟中，可以先進行MPEG-2格式的解碼，然後利用內插法(Interpolation)解交錯以得到視訊資料41。

在步驟203中，文字擷取模組104與插圖擷取模組105依據製作方針50來分析視訊資料41以取得文字部分801與插圖部分802，其能夠依據音訊分析演算法則501、字幕分析演算法則502、影像分析演算法則503以及場景/鏡頭變換分析演算法則504，針對視訊資料41的每一視訊畫面與內容(包含音訊內容)，進行分析搜尋並篩選取得符合製作方針50的文字部分801與插圖部分802，例如，若視訊資料41包括字幕串流則解讀視訊資料41之字幕串流以作為文字部分801；若視訊資料41未包括字幕串流則解讀視訊資料41之音訊，並利用語音分析進行語音與文字的轉換處理以作為文字部分801，並在與字幕串流或音訊對應之影像中擷取關鍵畫面作為插圖部分802，需注意者，本實施例可以擷取複數張關鍵畫面來作為插圖部分802。如圖3所示，原視訊資料40經過解碼後會得到視訊資料41，其包括複數張單張影像301(每秒25張或29.97張)，而經過依據製作方針50的分析搜尋後會從該等單張影像中擷取出關鍵畫面



五、發明說明 (14)

302 以作為插圖部分802。

步驟204係判斷是否已經完成視訊資料41中所有內容的分析比對，當未完成視訊資料41中所有內容的分析比對時，重複進行步驟203；當完成視訊資料41中所有內容的分析比對時，進行步驟205。

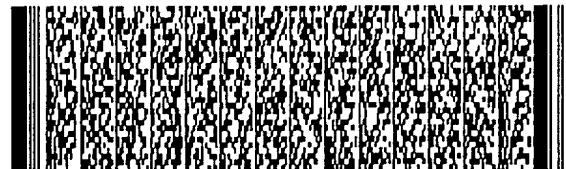
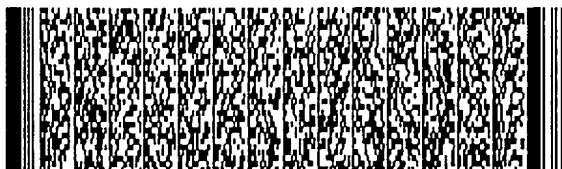
步驟205係判斷圖書80是否需要套用圖書格式70，當圖書80需要套用圖書格式70時，進行步驟206；當圖書80不需要套用圖書格式70時，進行步驟207。

在步驟206中，圖書格式選擇模組106提供使用者選擇所需之圖書格式70，圖書格式70包括各種具有圖片、影像、相片、繪畫或是繪圖之圖書樣板，例如，漫畫、繪本、畫冊、電子書等，以及各種佈置版面(Layout)。

在步驟207中，圖書產生模組107依據於步驟203中取得之文字部分801與插圖部分802，而且，當有進行步驟206時，套用步驟206中所提供的圖書格式70，並運用不同之濾鏡，如藝術家式濾鏡、素描濾鏡、邊緣濾鏡等，來處理插圖部分802，以得到所需之影像處理效果，再利用如調整大小，影像合成、製作圖框等影像處理功能得到符合圖書格式70之影像畫面，然後，將文字部分801與插圖部分802配合圖書格式70與字型、大小進行轉換處理，來產生圖書80。

步驟208係判斷使用者是否進行手動編輯圖書80，當使用者要進行手動編輯圖書80時，進行步驟209。

在步驟209中，使用者係利用編輯模組108來預覽

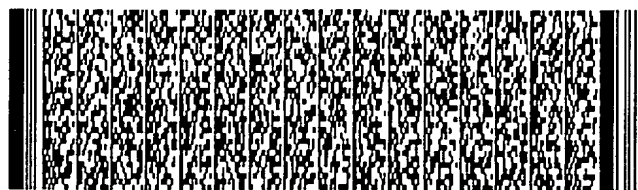


五、發明說明 (15)

(Preview)、修改(Refine)、修飾(Modify)圖書80之內容。例如，使用者可以針對圖書80之重要內容的文字部分加上底線，或是文字加粗等；或是使用者可以另外插入圖案等等。

綜上所述，由於依本發明較佳實施例之圖書製作系統與方法能夠分析視訊資料41，以針對視訊資料41之音訊資料411、字幕資料412及影像資料413，來整合視訊內容分析、文字辨識、聲音辨識等技術，所以能夠有效率地利用視訊資料來產生圖書文件。

以上所述僅為舉例性，而非為限制性者。任何未脫離本發明之精神與範疇，而對其進行之等效修改或變更，均應包含於後附之申請專利範圍中。



圖式簡單說明

【圖式簡單說明】

圖1為一示意圖，顯示依本發明較佳實施例之圖書製作系統之結構。

圖2為一流程圖，顯示依本發明較佳實施例之圖書製作方法的流程。

圖3為一示意圖，顯示於本發明較佳實施例之圖書製作方法中擷取關鍵畫面的示意圖。

【圖式符號說明】

1	圖書製作系統
101	視訊接收模組
102	解碼模組
103	製作方針選擇模組
104	文字擷取模組
105	插圖擷取模組
106	圖書格式選擇模組
107	圖書產生模組
108	編輯模組
2	圖書製作方法
201~209	依本發明較佳實施例之圖書製作方法的流音訊資料
301	單張影像
302	關鍵畫面
40	原視訊資料



圖式簡單說明

41	視訊資料
411	音訊資料
412	字幕資料
413	影像資料
50	製作方針
501	音訊分析演算法則
502	字幕分析演算法則
503	影像分析演算法則
5031	影像範例資料
5032	物體資料
504	場景/鏡頭變換分析演算法則
60	電腦設備
601	訊號源介面
602	記憶體
603	中央處理單元
604	輸入裝置
605	儲存裝置
70	圖書格式
80	圖書
801	文字部分
802	插圖部分



六、申請專利範圍

1、一種圖書製作系統，其係用來產生一圖書，該圖書包含一文字部分以及一插圖部分，該圖書製作系統包含：

一視訊接收模組，其接收一原視訊資料；

一解碼模組，其解碼該原視訊資料以取得一視訊資料；

一文字擷取模組，其依據一製作方針自該視訊資料中取得該文字部分；

一插圖擷取模組，其依據該製作方針自該視訊資料中擷取一關鍵畫面 (key frame) 以作為該插圖部分；以及

一圖書產生模組，其依據所取得之該文字部分與該插圖部分產生該圖書。

2、如申請專利範圍第1項所述之圖書製作系統，更包含：

一編輯模組，其於該圖書產生之後，接收一使用者之操作以編輯該圖書之內容。

3、如申請專利範圍第1項所述之圖書製作系統，更包含：

一圖書格式 (template) 選擇模組，其接收一使用者之選擇以提供至少一圖書格式，而該圖書產生模組係套用該圖書格式產生該圖書。

4、如申請專利範圍第1項所述之圖書製作系統，更包含：

一製作方針選擇模組，其接受一使用者之選擇以提供該製作方針。



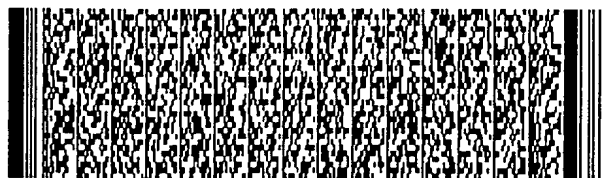
六、申請專利範圍

5、如申請專利範圍第1項所述之圖書製作系統，其中該製作方針包含一音訊（audio）分析演算法則，其分析該視訊資料中的一音訊資料，該文字擷取模組係依據該音訊分析演算法則擷取該音訊資料以取得該文字部分，而該插圖擷取模組係擷取與該音訊資料相對應之一影像資料以作為該插圖部分。

6、如申請專利範圍第1項所述之圖書製作系統，其中該製作方針包含一字幕（caption）分析演算法則，其係分析該視訊資料中的一字幕資料，該文字擷取模組係依據該字幕分析演算法則擷取該字幕資料以取得該文字部分，而該插圖擷取模組係擷取與該字幕資料相對應之一影像資料以作為該插圖部分。

7、如申請專利範圍第1項所述之圖書製作系統，其中該製作方針包含一影像分析演算法則，其係依據一影像範例分析該視訊資料中的一影像資料，該插圖擷取模組係依據該影像分析演算法則擷取該影像資料以取得該插圖部分，而該文字擷取模組係從與該影像資料相對應之該視訊資料中取得該文字部分。

8、如申請專利範圍第1項所述之圖書製作系統，其中該製作方針包含一影像分析演算法則，其係依據一物體分析該



六、申請專利範圍

視訊資料中的一影像資料，該插圖擷取模組係依據該影像分析演算法則擷取該影像資料以取得該插圖部分，而該文字擷取模組係從與該影像資料相對應之該視訊資料中取得該文字部分。

9、如申請專利範圍第1項所述之圖書製作系統，其中該製作方針包含一影像分析演算法則，其係分析該視訊資料中的一影像資料，該文字擷取模組係擷取該影像資料中的字幕以作為該文字部分，而該插圖擷取模組係擷取該影像資料以作為該插圖部分。

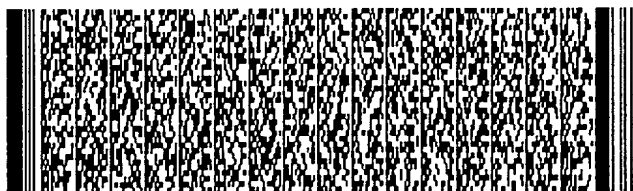
10、如申請專利範圍第1項所述之圖書製作系統，其中該製作方針包含一場景/鏡頭變換分析演算法則，其係分析該視訊資料中一影像資料的場景/鏡頭變換，該文字擷取模組與該插圖擷取模組係以該場景/鏡頭變換分析演算法則作為該文字部分與該插圖部分之選擇與分段的依據。

11、一種圖書製作方法，其係用來產生一圖書，該圖書包含一文字部分以及一插圖部分，該圖書製作方法包含：

一視訊接收步驟，其接收一原視訊資料；

一解碼步驟，其解碼該原視訊資料以取得一視訊資料；

一文字擷取步驟，其依據一製作方針自該視訊資料中取得該文字部分；



六、申請專利範圍

一 插圖擷取步驟，其依據該製作方針自該視訊資料中擷取一關鍵畫面以作為該插圖部分；以及

一圖書產生步驟，其依據所取得之該文字部分與該插圖部分產生該圖書。

12、如申請專利範圍第11項所述之圖書製作方法，更包含：

一編輯步驟，其於該圖書產生之後，接收一使用者之操作以編輯該圖書之內容。

13、如申請專利範圍第11項所述之圖書製作方法，更包含：

一圖書格式(template)選擇步驟，其接收一使用者之選擇以提供至少一圖書格式，而該圖書產生步驟係套用該圖書格式產生該圖書。

14、如申請專利範圍第11項所述之圖書製作方法，更包含：

一製作方針選擇步驟，其接受一使用者之選擇以提供該製作方針。

15、如申請專利範圍第11項所述之圖書製作方法，其中該製作方針包含一音訊分析演算法則，其分析該視訊資料中的一音訊資料，該文字擷取步驟係依據該音訊分析演算法



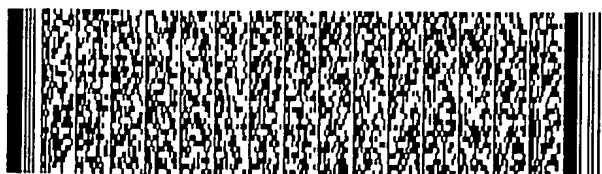
六、申請專利範圍

則擷取該音訊資料以取得該文字部分，而該插圖擷取步驟係擷取與該音訊資料相對應之一影像資料以作為該插圖部分。

16、如申請專利範圍第11項所述之圖書製作方法，其中該製作方針包含一字幕分析演算法則，其係分析該視訊資料中的一字幕資料，該文字擷取步驟係依據該字幕分析演算法則擷取該字幕資料以取得該文字部分，而該插圖擷取步驟係擷取與該字幕資料相對應之一影像資料以作為該插圖部分。

17、如申請專利範圍第11項所述之圖書製作方法，其中該製作方針係一影像分析演算法則，其係依據一影像範例分析該視訊資料中的一影像資料，該插圖擷取步驟係依據該影像分析演算法則擷取該影像資料以取得該插圖部分，而該文字擷取步驟係從與該影像資料相對應之該視訊資料中取得該文字部分。

18、如申請專利範圍第11項所述之圖書製作方法，其中該製作方針包含一影像分析演算法則，其係依據一物體分析該視訊資料中的一影像資料，該插圖擷取步驟係依據該影像分析演算法則擷取該影像資料以取得該插圖部分，而該文字擷取步驟係從與該影像資料相對應之該視訊資料中取得該文字部分。



六、申請專利範圍

19、如申請專利範圍第11項所述之圖書製作方法，其中該製作方針包含一影像分析演算法則，其係分析該視訊資料中的一影像資料，該文字擷取步驟係擷取該影像資料中的字幕以作為該文字部分，而該插圖擷取步驟係擷取該影像資料以作為該插圖部分。

20、如申請專利範圍第11項所述之圖書製作方法，其中該製作方針包含一場景/鏡頭變換分析演算法則，其係分析該視訊資料中一影像資料的場景/鏡頭變換，該文字擷取步驟與該插圖擷取步驟係以該場景/鏡頭變換分析演算法則作為該文字部分與該插圖部分之選擇與分段的依據。

21、一種記錄媒體，其係記錄用以使電腦達成一圖書製作方法之程式，該圖書製作方法係用來產生一圖書，該圖書包含一文字部分以及一插圖部分，該圖書製作方法包含：

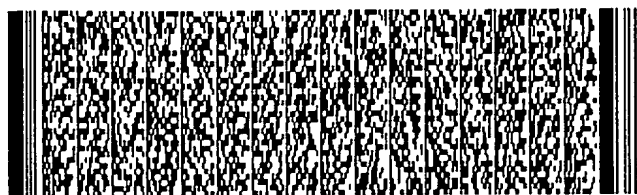
一視訊接收步驟，其接收一原視訊資料；

一解碼步驟，其解碼該原視訊資料以取得一視訊資料；

一文字擷取步驟，其依據一製作方針自該視訊資料中取得該文字部分；

一插圖擷取步驟，其依據該製作方針自該視訊資料中擷取一關鍵畫面以作為該插圖部分；以及

一圖書產生步驟，其依據所取得之該文字部分與該插



六、申請專利範圍

圖部分產生該圖書。

22、如申請專利範圍第21項所述之記錄媒體，其中該圖書製作方法更包含：

一編輯步驟，其於該圖書產生之後，接收一使用者之操作以編輯該圖書之內容。

23、如申請專利範圍第21項所述之記錄媒體，其中該圖書製作方法更包含：

一圖書格式(template)選擇步驟，其接收一使用者之選擇以提供至少一圖書格式，而該圖書產生步驟係套用該圖書格式產生該圖書。

24、如申請專利範圍第21項所述之記錄媒體，其中該圖書製作方法更包含：

一製作方針選擇步驟，其接受一使用者之選擇以提供該製作方針。

25、如申請專利範圍第21項所述之記錄媒體，其中該製作方針包含一音訊分析演算法則，其分析該視訊資料中的一音訊資料，該文字擷取步驟係依據該音訊分析演算法則擷取該音訊資料以取得該文字部分，而該插圖擷取步驟係擷取與該音訊資料相對應之一影像資料以作為該插圖部分。



六、申請專利範圍

26、如申請專利範圍第21項所述之記錄媒體，其中該製作方針包含一字幕分析演算法則，其係分析該視訊資料中的一字幕資料，該文字擷取步驟係依據該字幕分析演算法則擷取該字幕資料以取得該文字部分，而該插圖擷取步驟係擷取與該字幕資料相對應之一影像資料以作為該插圖部分。

27、如申請專利範圍第21項所述之記錄媒體，其中該製作方針係一影像分析演算法則，其係依據一影像範例分析該視訊資料中的一影像資料，該插圖擷取步驟係依據該影像分析演算法則擷取該影像資料以取得該插圖部分，而該文字擷取步驟係從與該影像資料相對應之該視訊資料中取得該文字部分。

28、如申請專利範圍第21項所述之記錄媒體，其中該製作方針包含一影像分析演算法則，其係依據一物體分析該視訊資料中的一影像資料，該插圖擷取步驟係依據該影像分析演算法則擷取該影像資料以取得該插圖部分，而該文字擷取步驟係從與該影像資料相對應之該視訊資料中取得該文字部分。

29、如申請專利範圍第21項所述之記錄媒體，其中該製作方針包含一影像分析演算法則，其係分析該視訊資料中的一影像資料，該文字擷取步驟係擷取該影像資料中的字幕



六、申請專利範圍

以作為該文字部分，而該插圖擷取步驟係擷取該影像資料以作為該插圖部分。

30、如申請專利範圍第21項所述之記錄媒體，其中該製作方針包含一場景/鏡頭變換分析演算法則，其係分析該視訊資料中一影像資料的場景/鏡頭變換，該文字擷取步驟與該插圖擷取步驟係以該場景/鏡頭變換分析演算法則作為該文字部分與該插圖部分之選擇與分段的依據。



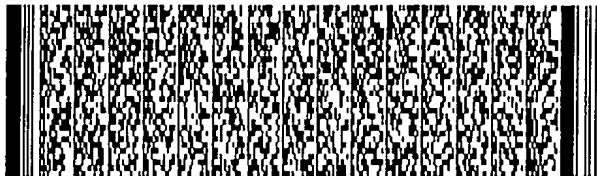
第 1/29 頁



第 2/29 頁



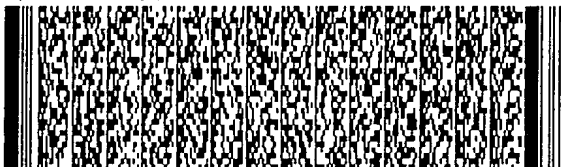
第 4/29 頁



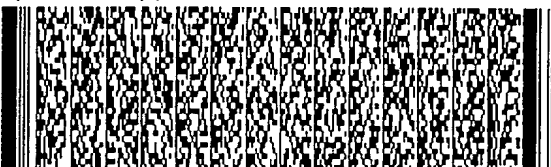
第 4/29 頁



第 5/29 頁



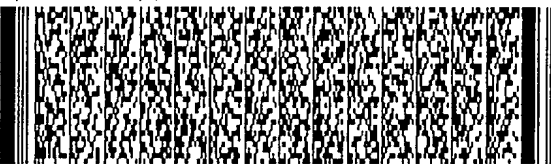
第 5/29 頁



第 6/29 頁



第 6/29 頁



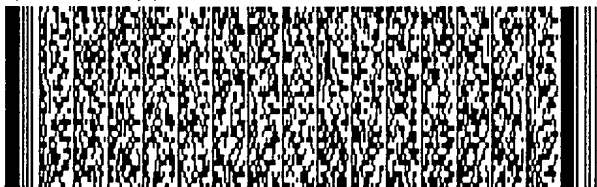
第 7/29 頁



第 7/29 頁



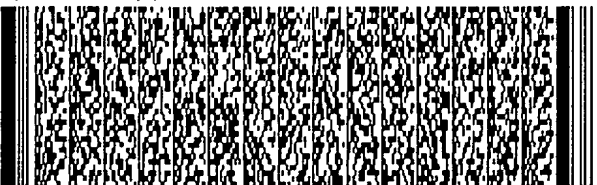
第 8/29 頁



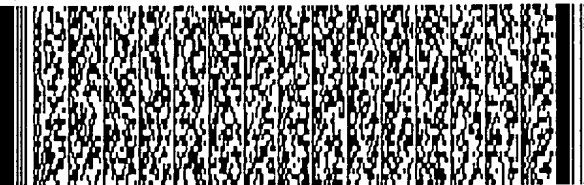
第 8/29 頁



第 9/29 頁



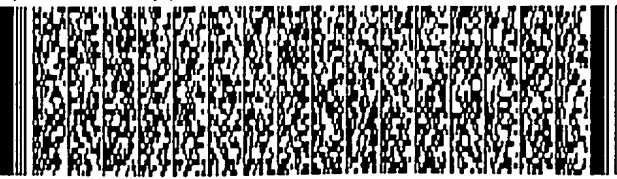
第 9/29 頁



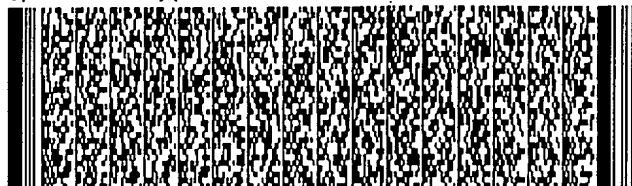
第 10/29 頁



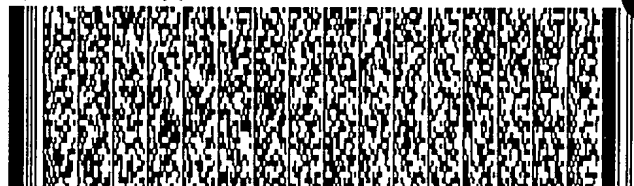
第 10/29 頁



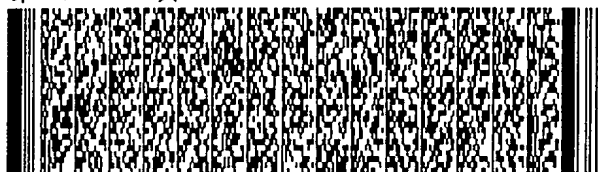
第 11/29 頁



第 11/29 頁



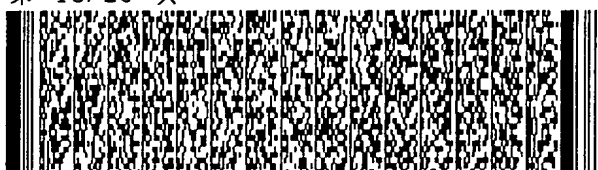
第 12/29 頁



第 12/29 頁



第 13/29 頁



第 13/29 頁



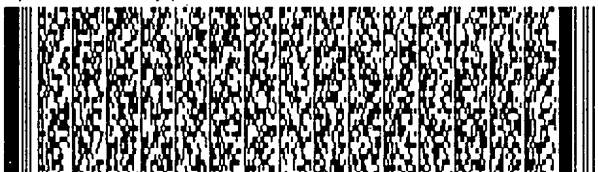
第 14/29 頁



第 14/29 頁



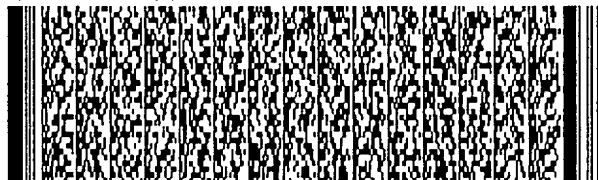
第 15/29 頁



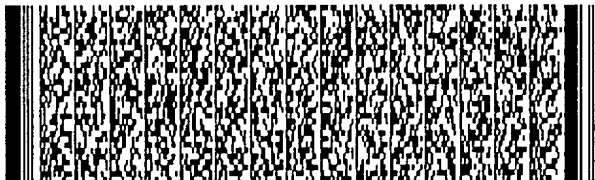
第 15/29 頁



第 16/29 頁



第 16/29 頁



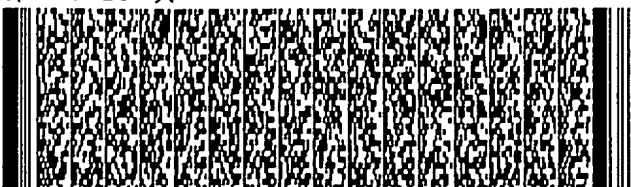
第 17/29 頁



第 17/29 頁



第 18/29 頁



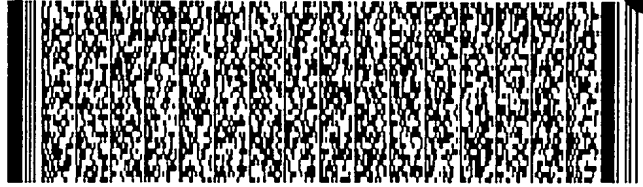
第 19/29 頁



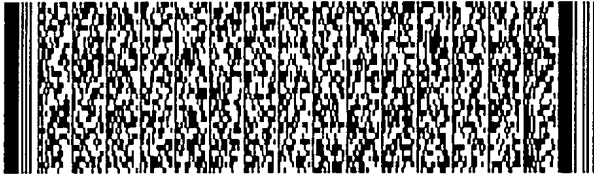
第 20/29 頁



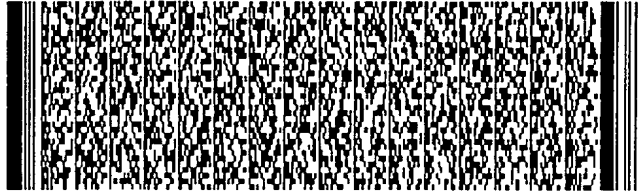
第 21/29 頁



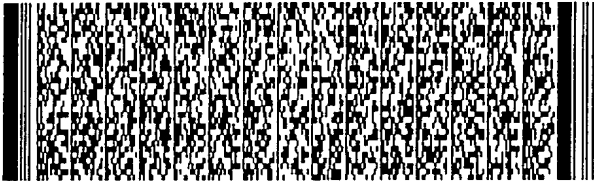
第 22/29 頁



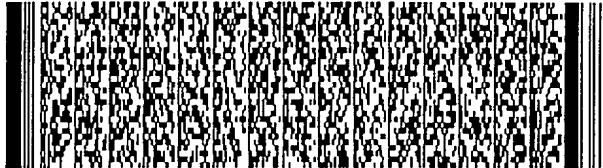
第 23/29 頁



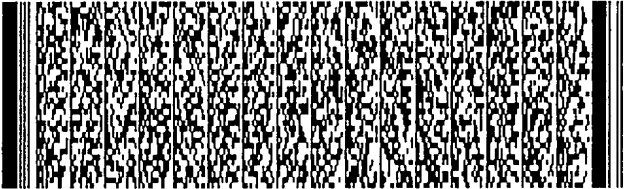
第 24/29 頁



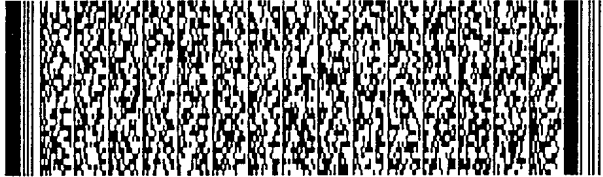
第 25/29 頁



第 26/29 頁



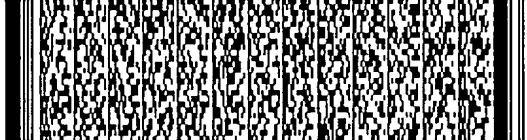
第 27/29 頁



第 28/29 頁



第 29/29 頁



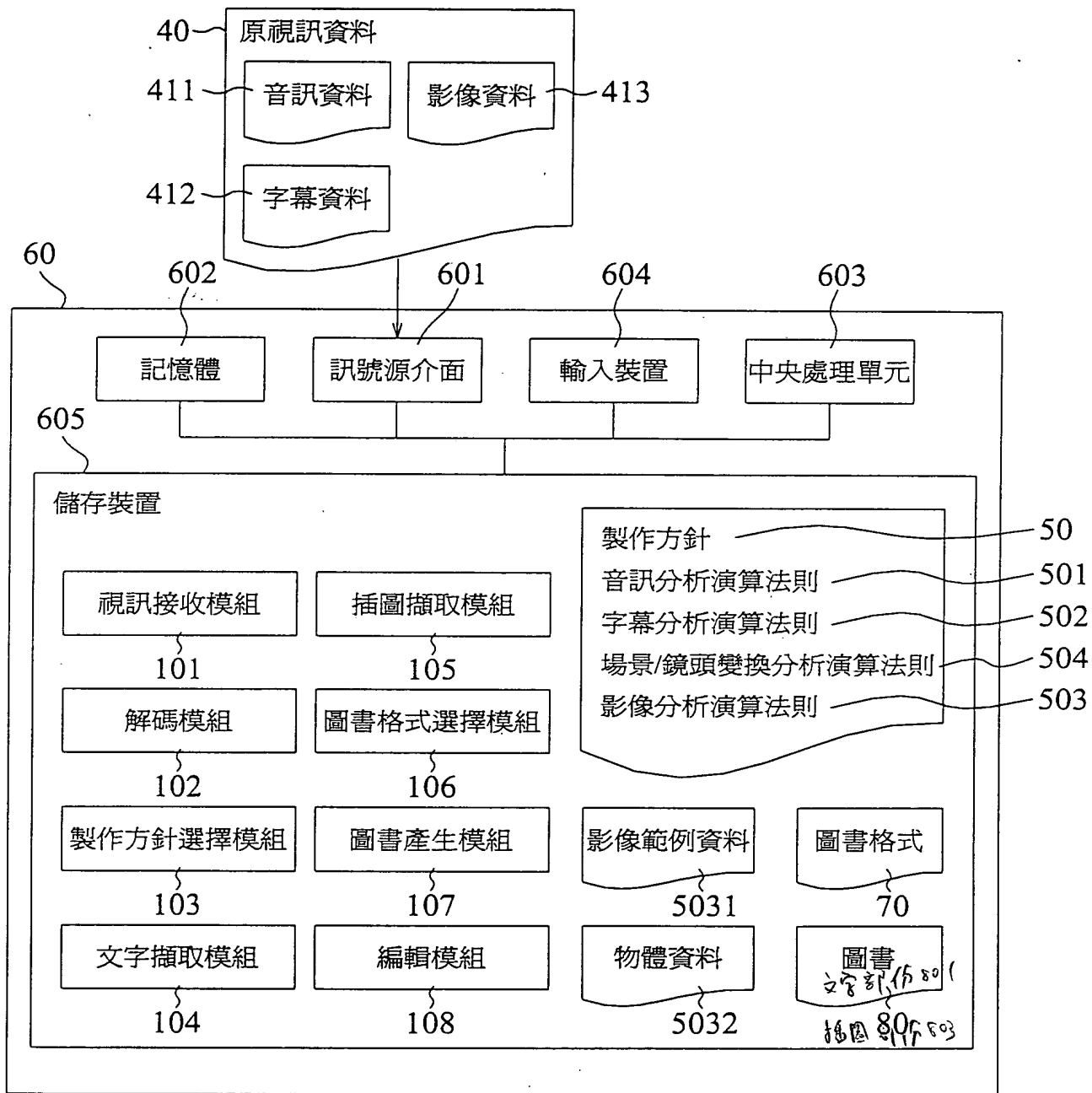


圖 1

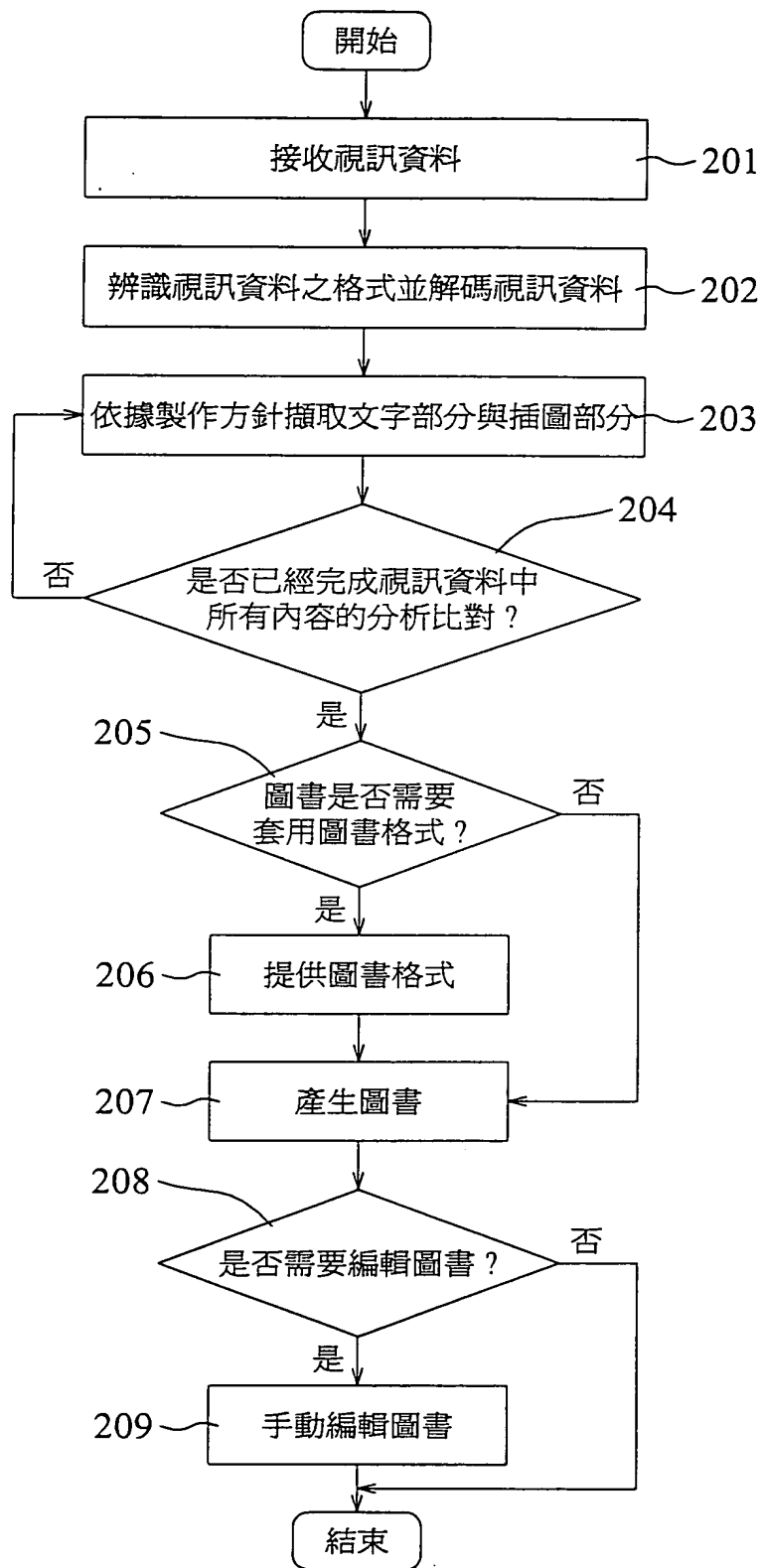


圖 2

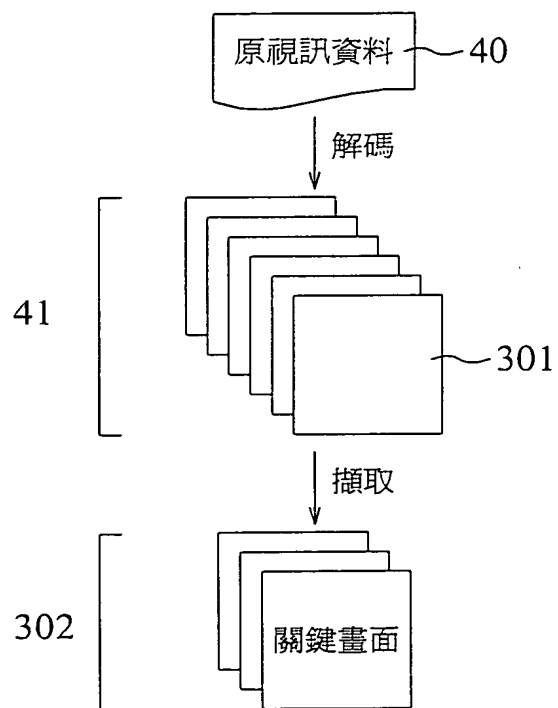


圖 3